

Internist 2008 · 49:681–687
 DOI 10.1007/s00108-008-2126-8
 Online publiziert: 10. April 2008
 © Springer Medizin Verlag 2008

Schwerpunktherausgeber
 M. Battagay, Basel
 J.F. Riemann, Ludwigshafen

H.C. Bucher

Basel Institute for Clinical Epidemiology & Biostatistics, Universitätsspital Basel

Studien mit Surrogatendpunkten

Nutzen und Grenzen in der klinischen Entscheidungsfindung

Behandlungsentscheidungen sollten sich möglichst auf die Ergebnisse von randomisierten kontrollierten Studien abstützen, welche auf patientenrelevanten Endpunkten wie z. B. verhinderte Krankheitsereignisse oder Tod basieren. Die Durchführung solcher Studien ist jedoch aufwändig und teuer, weshalb oft Studien mit Surrogatendpunkten durchgeführt werden. Diese Arbeit entwickelt Definitionen und Kriterien, anhand derer Kliniker die Ergebnisse von Surrogatendpunktstudien im Rahmen der klinischen Entscheidungsfindung kritisch gewichten lernen.

Das klinische Problem

Ein Pharmareferent stellt Ihnen Sitagliptin, ein Präparat einer neuen Substanzklasse von Dipeptidylpeptidase-IV-Inhibitoren zur Behandlung von Typ-2-Diabetikern vor, welche unter Metformin unzureichend eingestellt sind. Er unterstreicht, dass in einer randomisierten Multicenterstudie nach 30 Wochen bei 22,1% der Patienten unter Sitagliptin gegenüber 3,3% in der Placebogruppe das HbA_{1c} unter 7% lag (Differenz von HbA_{1c} zwischen den Gruppen –1,0% zugunsten von Sitagliptin; [22]). Er betont das niedrigere Hypoglykämierisiko von Sitagliptin in der Kombinationstherapie im Vergleich mit anderen Substanzen. Das Präparat sei eine echte Alternative gegenüber den „Problemmedikamenten aus der Klasse der Thiazolidinedionen“.

Wie glaubhaft sind die Ergebnisse dieser Studie? Ist Sitagliptin – ein Präparat, für welches Daten zur Verbesserung von bekannten klinischen Komplikation bei Typ-2-Diabetes fehlen, – eine echte Option für Ihre unzureichend eingestellten Typ-2-Diabetiker?

Was ist ein Surrogatendpunkt?

Behandlungsentscheidungen in der Klinik sollten sich im Idealfall auf die Ergebnisse von randomisierten kontrollierten Studien abstützen, welche den Nutzen einer Intervention anhand von Endpunkten untersuchen, die für Patienten relevant sind. Patientenrelevante Endpunkte sind zum Beispiel Lebensqualität, Herzinfarkt oder Tod. Da oft große Patientenzahlen und eine lange Beobachtungszeit benötigt werden, ist die Durchführung solcher Studien jedoch aufwändig. Aus diesem Grund werden von Forschern und der pharmazeutischen Industrie oft Studien durchgeführt, bei welchen patientenrelevante Endpunkte durch sog. Surrogatendpunkte ersetzt werden. Dies erlaubt, die Patientenzahl und Studiendauer zu reduzieren, und erscheint als eine Lösung für das Dilemma [1].

Surrogatendpunkte sind Ersatzmessparameter für die Lebensqualität, den organischen Funktionszustand oder die Überlebenswahrscheinlichkeit eines Patienten. Sie umfassen physiologische Variablen wie Albuminausscheidung im Urin als Surrogatendpunkt für dialysepflichtige Niereninsuffizienz, Knochendichte

als Surrogatendpunkt für osteoporosebedingte Frakturen, Blutdruck als Surrogatendpunkt für Hirnschlag und CD4-Zellzahl als Surrogatendpunkt für AIDS oder AIDS-assoziierte Mortalität [31]. Surrogatendpunkte werden manchmal auch als Biomarker bezeichnet.

Die Substitution eines patientenrelevanten Endpunkts mit einem Surrogatendpunkt ist v. a. dann sinnvoll, wenn der Surrogatmarker einfach, genügend früh im Krankheitsverlauf, mit hoher Präzision und ohne Beeinträchtigung durch Störfaktoren (z. B. beim Vorliegen von kompetitiven Risiken) gemessen werden kann. Ein nützlicher Surrogatmarker sollte nicht nur mit dem patientenrelevanten Endpunkt korreliert sein, sondern muss auch im höchst möglichen Ausmaß den therapeutischen Effekt auf den patientenrelevanten Endpunkt widerspiegeln.

Nützlich oder irreführend?

Surrogatendpunkte sind unerlässlich in der Evaluation und Nutzenbemessung von neuen Medikamenten in Phase-II- und frühen Phase-III-Studien. In vielen Ländern erhält die pharmazeutische Industrie die Zulassung eines neuen Medikaments basierend auf Daten von Surrogatendpunktstudien. Diese Zulassungspolitik spiegelt das Entscheidungs dilemma wider, in dem Zulassungsbehörden im Spannungsfeld von öffentlichem Interesse nach neuen, besseren und sichereren Medikamenten und dem Druck der pharmazeutischen Industrie stehen, ihre Pro-

dukte möglichst rasch auf den Markt zu bringen.

Sich auf Evidenz von Surrogatmarkerstudien zu verlassen, kann sinnvoll oder auch schädlich sein. Die Entscheidung der amerikanischen Zulassungsbehörde (Food and Drug Administration), neue antiretrovirale Medikamente aufgrund ihres Wirkungseffekts auf die Surrogatendpunkte CD4-Zellzahl und HI-Viruszelllast zuzulassen, war richtungsweisend und rettete Tausende von Patienten vor AIDS oder Tod. Nachfolgende patientenrelevante Endpunktstudien zeigten, dass eine antiretrovirale Therapie, z. B. mit einem Protease-Inhibitor, die HIV-assoziierte Morbidität und Mortalität senkt [11].

Andererseits kann das Vertrauen auf Surrogatendpunktstudien irreführend sein. Flosequinan, Milrinon, und Ibopamin sind alles Substanzen, welche wichtige hämodynamische Parameter bei Patienten mit Herzinsuffizienz verbessern. Nachfolgende randomisierte kontrollierte Studien zeigten jedoch, dass jede dieser Substanzen im Vergleich zu Placebo zu einer Exzessmortalität führte.

Wie können Kliniker zwischen einem glaubwürdigen und einem ungültigen Surrogatmarker unterscheiden? Anhand von klinischen Beispielen soll gezeigt werden, aufgrund welcher Kriterien die Gültigkeit von Surrogatmarkerstudien beurteilt werden kann ([2]; Infobox 1). Hierbei zeigt sich, dass die Beurteilung komplex ist und eine systematische Evaluation der Beziehung von Surrogat- und patientenrelevantem Endpunkt notwendig macht und die Aufarbeitung der Evidenz von allen randomisierten kontrollierten Studien einer Substanz mit patientenre-

levanten Endpunkten beinhaltet. Kliniker haben nicht die Zeit, sich dieses Wissen zu erarbeiten. Anhand von Beispielen sollen Kriterien zur Beurteilung der Gültigkeit von Surrogatmarkerstudien vorgestellt werden, welche Klinikern ermöglichen, die Argumente von Experten – oder der pharmazeutischen Industrie – zur Verschreibung von Medikamenten basierend auf Evidenz von Surrogatmarkerstudien kritisch zu evaluieren.

Beurteilung von Surrogatendpunktstudien

Sind die Ergebnisse von Surrogatmarkerstudien glaubhaft?

Assoziation zwischen Surrogat- und patientenrelevantem Endpunkt

Damit ein Surrogatendpunkt ein glaubwürdiger Ersatz für einen patientenrelevanten Endpunkt ist, muss eine hohe Assoziation zwischen dem Surrogatendpunkt und dem patientenrelevanten Endpunkt bestehen. Je höher das Assoziationsmaß, desto wahrscheinlicher ist der kausale Zusammenhang zwischen Surrogat- und patientenrelevantem Endpunkt. D. h. der Surrogatmarker muss im kausalen pathophysiologischen Prozess bedeutsam sein [1]. Zum Beispiel muss LDL-Cholesterin eine Ursache von ischämischen kardiovaskulären Ereignissen sein, um ein glaubhafter Surrogatmarker zu sein. Zudem muss eine Änderung des Surrogatmarkers (z. B. des LDL-Cholesterins) einen relevanten Einfluss auf patientenrelevante Endpunkte haben. Wenn eine Intervention unabhängig vom Surrogatendpunkt Morbidität oder Mortali-

tät positiv oder negativ (wie beispielsweise Fibrate zur Prävention von Myokardinfarkten; [29]) beeinflusst, ist die Glaubwürdigkeit des Surrogatendpunkts in Frage gestellt.

Die Glaubwürdigkeit eines Surrogatmarkers hängt zudem davon ab, dass in verschiedenen Studien ein konsistenter und unabhängiger Zusammenhang zwischen dem Surrogatmarker und patientenrelevanten Endpunkten besteht. Beispielsweise bestätigen große Kohortenstudien unabhängig von anderen Risikofaktoren wie Alter, Geschlecht, Blutdruck und Rauchen den Zusammenhang der Höhe des Cholesterinspiegels und dem Myokardinfarktrisiko [35].

Zurück zu unserem Beispiel: Evidenz aus randomisierten kontrollierten Studien bei Typ-1- und -2-Diabetes zeigt, dass eine bessere Kontrolle des Blutzuckers mikrovaskuläre Komplikationen reduziert [4, 32]. Große Kohortenstudien zeigen einen Zusammenhang zwischen dem Ausmaß der Blutzuckerkontrolle und dem Risiko von makrovaskulären Komplikationen [13, 14]. In einer Metaanalyse von 13 Kohortenstudien (unter Einschluss der UKPD-Studie) nahm pro Prozent Zunahme des HbA_{1c} das relative Risiko eines kardiovaskulären Ereignisses zu (RR 1,18; 95% CI 1,10–1,26; [26]). Damit ist das erste Kriterium für einen glaubwürdigen Surrogatendpunkt erfüllt.

Übereinstimmung mit randomisierten Studien verschiedener Medikamentenklassen

Pathophysiologische Studien sowie nicht-experimentelle Beobachtungsstudien sind unzureichend, um den kausalen Zusammenhang zwischen Surrogatendpunkt und patientenrelevanten Endpunkten zu dokumentieren. Antiarrhythmika der Klasse I zeigten in elektrophysiologischen Untersuchungen ausgezeichnete Eigenschaften, komplexe ventrikuläre Rhythmusstörungen zu unterdrücken und wurden deshalb häufig bei Hochrisikopatienten eingesetzt. Als – mit beträchtlicher Verzögerung – Encainid und Flecainid im Rahmen einer randomisierten kontrollierten Studie überprüft wurden, musste die Studie aufgrund einer massiven Exzessmortalität beider Substanzen gegen-

Infobox 1

Kriterien zur Beurteilung von Surrogatendpunktstudien

Sind die Ergebnisse von Surrogatmarkerstudien glaubhaft?

- Besteht eine starke, unabhängige und konsistente Assoziation zwischen dem Surrogatendpunkt und dem patientenrelevanten Endpunkt?
- Zeigen randomisierte Studien verschiedener Medikamentenklassen, dass die Veränderung des Surrogatendpunkts konsistent patientenrelevante Endpunkte verbessert?
- Zeigen randomisierte Studien der gleichen Medikamentenklassen, dass die Veränderung des Surrogatendpunkts konsistent patientenrelevante Endpunkte verbessert?

Wie lauten die Ergebnisse?

- Wie groß, präzise und anhaltend ist der Behandlungseffekt?

Wie können die Ergebnisse in die Patientenversorgung einfließen?

- Stehen die wahrscheinlichen Behandlungsvorteile im Verhältnis zu möglichen Risiken und den Kosten?

über Placebo vorzeitig abgebrochen werden [7].

Die Behandlung der Herzinsuffizienz bietet ein weiteres instruktives Beispiel. Kontrollierte Studien bei Patienten mit Herzinsuffizienz zeigen, dass ACE-Hemmer sowohl die Belastungskapazität verbessern [6, 23] als auch zu einer Mortalitätsreduktion führen [8] und damit die Belastungskapazität einen für Kliniker nützlichen Surrogatendpunkt darstellt. Milrinon [5] und Epoprostenol [30] zeigten in der Behandlung der Herzinsuffizienz eine deutliche Verbesserung der Belastungskapazität, waren jedoch in randomisierten kontrollierten Studien mit einer erhöhten Mortalität assoziiert [3, 20]. Ipopamin zeigte eine Verbesserung von gleich 3 Surrogatendpunkten (Auswurfraction, Variabilität der Herzrate und Markern autonomer Funktion; [33]), doch in der Überprüfung im Rahmen einer randomisierten Studie resultierte insbesondere bei Patienten mit Herzinsuffizienz wiederum eine Exzessmortalität [12].

Über Jahre wurde Natriumfluorid aufgrund des günstigen Einflusses auf die Knochendichte in der Behandlung der Osteoporose eingesetzt. Im Rahmen einer randomisierten kontrollierten Studie zeigten jedoch mit Natriumfluorid behandelte Frauen trotz Zunahme der Knochendichte aufgrund einer erhöhten Knochensprödigkeit vermehrte Frakturen [24]. Der ungünstige Einfluss von Natriumfluorid auf die Frakturrate wurde später in einer Metaanalyse von randomisierten kontrollierten Studien bestätigt [10].

Es gibt jedoch auch Beispiele von günstigen Surrogatmarkern. Erste randomisierte kontrollierte Studien mit Protease-Inhibitoren bei HIV-Infizierten zeigten, dass mit vollständiger Unterdrückung der HI-Viruslast und dem Anstieg der CD4-Zellzahl AIDS und die HIV-assozierte Mortalität relevant gesenkt werden können [11]. Nachfolgende große Kohortenstudien zeigten, dass neuere antiretrovirale Substanzen zu einer weiteren bedeutsamen Mortalitätsreduktion führen [28]. Diese Ergebnisse ermutigen uns in der Schlussfolgerung, dass neue Substanzen wie z. B. Darunavir, ein neuer Protease-Inhibitor [16], oder Raltegravir, ein Integraseblocker [9], welche bei Patienten mit fortgeschrittener HIV-Infektion und

Hier steht eine Anzeige.



Tab. 1 Beispiele von Surrogatendpunkten und Kriterien zur Beurteilung der Glaubwürdigkeit von Studienergebnissen mit Surrogatendpunkten

Art der Intervention	Surrogatendpunkt	Patienten-relevanter Endpunkt	Kriterium 1	Kriterium 2	Kriterium 3
			Besteht eine starke, unabhängige und konsistente Assoziation zwischen dem Surrogatendpunkt und dem patientenrelevanten Endpunkt?	Zeigen randomisierte Studien <i>verschiedener</i> Medikamentenklassen, dass die Veränderung des Surrogatendpunkts konsistent patientenrelevante Endpunkte verbessert?	Zeigen randomisierte Studien der <i>gleichen</i> Medikamentenklassen, dass die Veränderung des Surrogatendpunkts konsistent patientenrelevante Endpunkte verbessert?
DPP-IV-Hemmer ^a Sitagliptin	HbA _{1c}	Kardiovaskuläre Ereignisse	Ja [13, 26]	Nein [15, 25, 27]	Nein [22]
Integraseblocker Raltegravir ^b	HI-Viruslast CD4-Zellzahl	AIDS oder Tod AIDS oder Tod	Ja [17] Ja [34]	Ja [11, 18] Ja [11]	Nein [9] Nein [9]
Protease-Inhibitor Darunavir ^b	HI-Viruslast CD4-Zellzahl	AIDS oder Tod AIDS oder Tod	Ja [17] Ja [34]	Ja [18] Ja [18]	Ja [18] Ja [18]
Lipidsenker Rosuvastatin	LDL-Cholesterinsenkung	Myokardinfarkt, Tod	Ja [35]	Nein [29]	Ja [29]

^a Dipeptidylpeptidase-IV-Hemmer, ^b In Kombinationstherapie mit 2 Reverse-Transkriptase-Hemmern.

multiplen Resistenzen zu einer eindrucksvollen Verbesserung der Surrogatendpunkte führten, auch zu einer relevanten Verbesserung der klinischen Prognose bei diesen Patientengruppen führen werden.

Zurück zum Beispiel: Neben Insulin stehen heute 6 Medikamentenklassen zur oralen Behandlung des Typ-2-Diabetes zur Verfügung [19]. Studien mit Ergebnissen zu patientenrelevanten Endpunkten liegen jedoch nur für Metformin und die Thiozolidnedione Rosiglitazon und Pioglitazon vor. In einer Metaanalyse reduzierte eine intensivisierte Blutzuckerkontrolle unter Monotherapie mit Metformin das Risiko der diabetesassoziierten Mortalität, von letalen Myokardinfarkten sowie die Gesamtmortalität [25]. Die Kombination von Metformin und Sulfonylharnstoffen führte jedoch in der UKPD-Studie zu einer 96%igen Exzessmortalität [32]. In Metaanalysen von randomisierten kontrollierten Studien zeigte Rosiglitazon im Vergleich zu anderen Substanzen keine bessere Senkung des HbA_{1c}, war jedoch in einer Metaanalyse mit einem erhöhten Risiko von Myokardinfarkten und Herzinsuffizienz assoziiert [27]. Eine weitere Metaanalyse fand für Rosiglitazon und Pioglitazon ebenfalls ein erhöhtes Risiko von Herzinsuffizienz und legt einen Klasseneffekt nahe, konnte jedoch aufgrund anderer

Einschlusskriterien von Studien eine erhöhte kardiovaskuläre Mortalität dieser Substanzen nicht erhärten [15]. Der Nutzen oraler Antidiabetika für patientenrelevante Endpunkte ist mit Ausnahme von Metformin unklar. Kriterium 2 für die Validität von HbA_{1c} als Surrogatendpunkt ist somit nicht erfüllt.

Übereinstimmung mit randomisierten Studien der gleichen Medikamentenklasse

Surrogatendpunktstudien können für Kliniker glaubwürdiger sein, wenn für Substanzen der gleichen Klasse randomisierte kontrollierte Studien vorliegen, welche eine Verbesserung von Surrogat- und patientenrelevanten Endpunkten zeigen. Beispielsweise zeigen mehrere große randomisierte kontrollierte Studien, dass Statine in der Primär- und Sekundärprävention kardiovaskuläre Ereignisse reduzieren [29]. Deshalb ist nahe liegend anzunehmen, dass Rosuvastatin, ein noch potenteres Präparat zur Senkung des LDL-Cholesterins, einen ebensolchen Effekt auf patientenrelevante Endpunkte hat. Hierbei gilt es jedoch zu bedenken, dass Cerivastatin – ein Präparat, welches nur aufgrund von Daten aus Surrogatmarkerstudien die Zulassung erhielt – wegen eines 10-fach erhöhten Risikos von Rhabdomyolysen gegenüber herkömmlichen

Statinen vom Markt genommen werden musste.

Das Vertrauen auf Surrogatendpunktstudien ist immer mit offenen Fragen zu Langzeitverträglichkeit und Toxizität verbunden.

Unsere Annahmen für dieses Kriterium können durch 2 Mechanismen außer Kraft gesetzt werden. So kann ein Präparat einen nützlichen Effekt außerhalb des über den Surrogatmarker vermittelten Effekts haben. Beispielsweise können Statine sich in ihren antiinflammatorischen Effekten, welche für die Endothelfunktion wichtig sind, unterscheiden. Zweitens können Substanzen schädliche Effekte haben, welche nicht über den Surrogatmarkermechanismus vermittelt sind. Beispielsweise ist die Zunahme der nicht kardiovaskulären Mortalität unter Fibraten einem anderen Mechanismus zuzuschreiben und nicht durch die Reduktion des LDL-Cholesterins und die damit assoziierte Reduktion von letalen und nicht-letalen Myokardinfarkten von Fibraten bedingt.

Unsere Kriterien werden durch die Definition des Klassenbegriffs einer Substanz tangiert. Es ist nahe liegend, dass ein Produzent geneigt ist, den Klassenbegriff für seine Substanz weit zu fas-

H.C. Bucher

Studien mit Surrogatendpunkten. Nutzen und Grenzen in der klinischen Entscheidungsfindung

Zusammenfassung

Therapieentscheidungen sollten im Idealfall auf die Ergebnisse von randomisierten kontrollierten Studien mit patientenrelevanten Endpunkten gestützt sein. Patientenrelevante Endpunkte sind zum Beispiel Lebensqualität, Herzinfarkt oder Tod. Die Durchführung solcher Studien ist jedoch aufwändig, da oft große Patientenzahlen und eine lange Beobachtungszeit benötigt werden. Deshalb werden oft Studien durchgeführt, bei welchen patientenrelevante Endpunkte durch Surrogatendpunkte ersetzt werden, um Patientenzahl und Beobachtungszeit zu verringern. Surrogatendpunkte sind Ersatzmessparameter für die Lebensqualität, den organischen Funktionszustand oder die Überlebenswahrscheinlichkeit eines Patienten. Die

Zulassung von neuen Substanzen beruht oft auf den Ergebnissen von Surrogatendpunktstudien. Unerwartete Nebenwirkungen und der Rückzug von Substanzen, welche aufgrund solcher Studien zugelassen wurden, haben in jüngster Zeit eine Kontroverse über die Bedeutung von Surrogatendpunktstudien entfacht. Wir erläutern anhand von Beispielen Kriterien, mit deren Hilfe Kliniker die Ergebnisse von Surrogatendpunktstudien kritisch gewichten können.

Schlüsselwörter

Surrogatendpunkt · Patientenrelevanter Endpunkt · Randomisierte kontrollierte Studie · Evidenzbasierte Medizin

Surrogate endpoint trials: Benefit and pitfalls for clinical decision making

Abstract

Ideally clinicians should base their treatment decisions on results from randomised controlled trials which include patient-important outcomes, such as quality of life, prevented disease events or death. Conducting such trials often involves large sample sizes and extended follow-up periods. Therefore, researchers have aimed to conduct trials with surrogate endpoints by substituting patient-important outcomes in order to reduce sample size and observation time. Surrogate endpoints are outcomes that substitute for direct measures of how a patient feels, functions, or survives. In many countries drugs are approved based on data from surrogate endpoint trials. Recently, a controversy evolved

on the reliability of results generated from these trials driven by unanticipated side effects or severe toxicity leading to the withdrawal of drugs that were solely approved based on evidence from surrogate endpoint trials. We present some recent examples and criteria how clinicians can critically evaluate the validity of claims by experts or the pharmaceutical industry in regard to the expected patients' benefit from drugs approved by results from surrogate endpoint trials.

Keywords

Surrogate endpoints · Clinical endpoint · Randomised controlled trial · Evidence-based medicine

sen, wenn mehrere Präparate der gleichen Klasse eine positive und konsistente Assoziation zwischen Surrogatendpunkt und patientenrelevantem Endpunkt gezeigt haben (Beispiel β -Blocker für Patienten nach Myokardinfarkt). Bei Substanzen mit bekanntem Toxizitätspotenzial werden Produzenten eher dazu neigen, die pharmakologischen Unterschiede ihrer Substanz gegenüber Substanzen der gleichen Klasse zu unterstreichen (Beispiel Clofibrat gegenüber einem anderen Fibrat).

Anhand kontroverser Beispiele von Surrogatendpunktstudien illustriert **Tab. 1** unsere Kriterien zur Beurteilung der Glaubwürdigkeit.

In unserem klinischen Beispiel ist Sitagliptin die erste Substanz der DPP-IV-Inhibitorenklasse und erhielt die Zulassung in der EU im April 2007 aufgrund von Ergebnissen von Surrogatendpunktstudien. Ergebnisse zu klinischen Endpunktstudien fehlen für diese Medikamentenklasse gänzlich.

Wie lauten die Ergebnisse?

Wie groß, präzise und anhaltend ist der Behandlungseffekt?

Unser Vertrauen in die Ergebnisse von Surrogatmarkerstudien kann erhöht werden, wenn der Surrogatmarker in hohem, zeitlich ausreichend langem und klinisch relevantem Ausmaß sowie mit genügender Präzision (enge 95%-Vertrauensintervalle um das Effektmaß der Surrogatmarkerveränderung) durch eine neue Substanz beeinflusst wird. Kleine, klinisch wenig relevante Veränderungen des Surrogatmarkers von ungewisser Dauer würden unser Vertrauen in einen Surrogatendpunkt schmälern.

Zurück zum Beispiel: In der Studie von Raz et al. [22] ergab sich bei Patienten, die mit Metformin behandelt waren, unter zusätzlichen Behandlung von Sitagliptin gegenüber Placebo nach 30 Wochen eine Differenz des HbA_{1c} von –1,0% (–1,4% bis –0,6%). Das Vertrauensintervall ist weit und schließt einen klinisch wenig relevanten Effekt ein. Es wurden 190 Patienten randomisiert und der Verlust an Patienten im Studienverlauf lag bei rund 10%. Insgesamt wurden unter Sitagliptin 15,6% Laborabnormalitäten – die nicht im Detail aufgeführt sind – aufgeführt, gegen-

über 4,3% in der Placebogruppe. In beiden Gruppen trat jeweils eine Hypoglykämie auf.

Wie können die Ergebnisse in die Patientenversorgung einfließen?

Verhältnis von Behandlungsvorteilen zu Risiken und Kosten

Die nachfolgenden Fragen muss sich ein Kliniker bei jeder Intervention stellen:

- Sind die Ergebnisse einer Studie auf einen gegebenen Patienten übertragbar?
- Wurden alle patientenrelevanten Endpunkte berücksichtigt?
- Überwiegt der erwartete Nutzen die möglichen Risiken?
- Ist die Maßnahme kosteneffektiv?

Die Abschätzung des zu erwartenden Nutzens für den Patienten ist bei Surrogatendpunktstudien eine besondere Herausforderung und verlangt die Übertragung eines zu erwartenden Nutzens der Modifikation des Surrogatmarkers auf den klinisch relevanten Endpunkt durch Extrapolation von anderen Studien.

In unserem Beispiel folgern wir, dass die Dokumentation des Nutzens von Sitagliptin auf relativ kleinen randomisierten Studien von mit Metformin vorbehandelten und ungenügend eingestellten Typ-2-Diabetikern beruht. Daten zur Langzeitsicherheit und dem Effekt von Sitagliptin auf patientenrelevante Endpunkte fehlen. Durch Herleitung von anderen Studien ist eine Differenz des HbA_{1c} von 1% mit einer 18%igen Risikoerhöhung von diabetesassoziierten patientenrelevanten Endpunkten verbunden. Die Verbesserung des HbA_{1c} durch Sitagliptin ist jedoch in der Studie von Raz et al. [22] mit großer Unsicherheit behaftet: HbA_{1c} erfüllt bloß 1 von 3 Kriterien für die Glaubwürdigkeit eines Surrogatendpunkts.

Aus diesem Grund entscheiden Sie sich aufgrund der großen Unsicherheit in der Evidenzlage, Sitagliptin bei Ihren ungenügend kontrollierten Patienten äußerst zurückhaltend einzusetzen und aggressiven Maßnahmen zur Verbesserung des Lebensstils zunächst den Vorzug zu geben [21].

Fazit für die Praxis

Therapieentscheidungen, welche sich auf Ergebnisse von Surrogatendpunktstudien stützen, implizieren gewichtige Annahmen bezüglich eines Nutzens für einen für den Patienten relevanten Endpunkt. Unse- re Kriterien können Klinikern helfen, sich dieser Annahmen explizit bewusst zu werden. Auch wenn ein Surrogatmarker al- le diese Kriterien erfüllt, können Kliniker und Patienten dennoch in die Irre geführt werden. Aus diesem Grund ist es entschei- dend, dass sich Kliniker über die gewich- tigen Probleme von Surrogatmarkerstu- dien bewusst sind.

Leider besteht eine nicht unerheblich lan- ge Liste von Beispielen – einige wurden hier aufgeführt – bei denen Tausende von Patienten durch Verlass auf Studienergeb- nisse von Surrogatendpunktstudien zu Schaden gekommen sind. Kliniker mit ei- ner eher konservativen Einstellung, wel- che die Ergebnisse von randomisierten kontrollierten Studien mit patientenrele- vanten Endpunkten abwarten, bevor sie ein neues Präparat verordnen, sind oft gut beraten, insbesondere wenn es Behand- lungsalternativen gibt. Ergebnisse von Surrogatmarkerstudien sollten möglichst nur dann Eingang in ein Behandlungskon- zept finden, wenn die Krankheit lebensbe- drohlich ist, ein sehr hoher Leidensdruck des Patienten besteht, die Effektgröße der Intervention auf den Surrogatendpunkt groß, präzise und klinisch relevant ist und keine gute andere Behandlungsalterna- tive vorliegt.

Korrespondenzadresse

Prof. Dr. H.C. Bucher

Basel Institute for Clinical Epidemiology & Bio-
statistics, Universitätsspital Basel
4031 Basel
Schweiz
bucherh@uhbs.ch

Interessenkonflikt. Der korrespondierende Autor gibt an, dass kein Interessenkonflikt besteht.

Literatur

1. Biomarkers Definitions Working Group (2001) Bio- markers and surrogate endpoints: preferred defi- nitions and conceptual framework. Clin Pharmacol Ther 69: 89–95

2. Bucher HC, Guyatt GH, Cook DJ et al. (1999) Users' guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring the effect of an intervention on surrogate end points. Evidence-Based Medicine Working Group. JAMA 282: 771–778
3. Califf RM, Adams KF, McKenna WJ et al. (1997) A randomized controlled trial of epoprostenol therapy for severe congestive heart failure: The Flolan International Randomized Survival Trial (FIRST). Am Heart J 134: 44–54
4. Diabetes Control and Complications Trial Research Group (1993) The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 329: 977–986
5. DiBianco R, Shabetai R, Kostuk W et al. (1989) A comparison of oral milrinone, digoxin, and their combination in the treatment of patients with chronic heart failure. N Engl J Med 320: 677–683
6. Drexler H, Banhardt U, Meinertz T et al. (1989) Contrasting peripheral short-term and long-term effects of converting enzyme inhibition in patients with congestive heart failure. A double-blind, placebo-controlled trial. Circulation 79: 491–502
7. Echt DS, Liebson PR, Mitchell LB et al. (1991) Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. N Engl J Med 324: 781–788
8. Garg R, Yusuf S (1995) Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor Trials. JAMA 273: 1450–1456
9. Grinsztejn B, Nguyen BY, Katlama C et al. (2007) Safety and efficacy of the HIV-1 integrase inhibitor raltegravir (MK-0518) in treatment-experienced patients with multidrug-resistant virus: a phase II randomised controlled trial. Lancet 369: 1261–1269
10. Hagenauer D, Welch V, Shea B et al. (2000) Fluoride for the treatment of postmenopausal osteoporotic fractures: a meta-analysis. Osteoporos Int 11: 727–738
11. Hammer SM, Squires KE, Hughes MD et al. (1997) A controlled trial of two nucleoside analogues plus didanosine in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. Aids Clinical Trials Group 320 Study Team. N Engl J Med 337: 725–733
12. Hampton JR, Van Veldhuisen DJ, Kleber FX et al. (1997) Randomised study of effect of ibopamine on survival in patients with advanced severe heart failure. Second Prospective Randomised Study of Ibopamine on Mortality and Efficacy (PRIME II) Investigators. Lancet 349: 971–977
13. Khaw KT, Wareham N, Bingham S et al. (2004) Association of hemoglobin A1c with cardiovascular disease and mortality in adults: the European prospective investigation into cancer in Norfolk. Ann Intern Med 141: 413–420
14. Kuusisto J, Mykkanen L, Pyorala K, Laakso M (1994) NIDDM and its metabolic control predict coronary heart disease in elderly subjects. Diabetes 43: 960–967
15. Lago RM, Singh PP, Nesto RW (2007) Congestive heart failure and cardiovascular death in patients with prediabetes and type 2 diabetes given thiazolidinediones: a meta-analysis of randomised clinical trials. Lancet 370: 1129–1136

16. Madruga JV, Berger D, McMurchie M et al. (2007) Efficacy and safety of darunavir-ritonavir compared with that of lopinavir-ritonavir at 48 weeks in treatment-experienced, HIV-infected patients in TITAN: a randomised controlled phase III trial. *Lancet* 370: 49–58
17. Mellors JW, Kingsley LA, Rinaldo CR et al. (1995) Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. *Ann Intern Med* 122: 573–579
18. Montaner JS, Reiss P, Cooper D et al. (1998) A randomized, double-blind trial comparing combinations of nevirapine, didanosine, and zidovudine for HIV-infected patients: the INCAS Trial. Italy, the Netherlands, Canada and Australia Study. *JAMA* 279: 930–937
19. Nathan DM (2007) Finding new treatments for diabetes – how many, how fast... how good? *N Engl J Med* 356: 437–440
20. Packer M, Carver JR, Rodeheffer RJ et al. (1991) Effect of oral milrinone on mortality in severe chronic heart failure. The Promise Study Research Group. *N Engl J Med* 325: 1468–1475
21. Pi-Sunyer X, Blackburn G, Brancati FL et al. (2007) Reduction in weight and cardiovascular disease risk factors in individuals with type 2 diabetes: one-year results of the look AHEAD trial. *Diabetes Care* 30: 1374–1383
22. Raz I, Chen Y, Wu M et al. (2008) Efficacy and safety of sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes. *Curr Med Res Opin* 24: 537–550
23. Riegger GA (1991) Effects of quinapril on exercise tolerance in patients with mild to moderate heart failure. *Eur Heart J* 12: 705–711
24. Riggs BL, Hodgson SF, O'Fallon WM et al. (1990) Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med* 322: 802–809
25. Saenz A, Fernandez-Esteban I, Mataix A et al. (2005) Metformin monotherapy for type 2 diabetes mellitus. *Cochrane Database Syst Rev*: CD002966
26. Selvin E, Marinopoulos S, Berkenblit G et al. (2004) Meta-analysis: glycosylated hemoglobin and cardiovascular disease in diabetes mellitus. *Ann Intern Med* 141: 421–431
27. Singh S, Loke YK, Furberg CD (2007) Long-term risk of cardiovascular events with rosiglitazone: a meta-analysis. *JAMA* 298: 1189–1195
28. Sterne JA, Hernan MA, Ledergerber B et al. (2005) Long-term effectiveness of potent antiretroviral therapy in preventing AIDS and death: a prospective cohort study. *Lancet* 366: 378–384
29. Studer M, Briel M, Leimenstoll B et al. (2005) Effect of different antilipidemic agents and diets on mortality: a systematic review. *Arch Intern Med* 165: 725–730
30. Sueta CA, Gheorghiade M, Adams KF et al. (1995) Safety and efficacy of epoprostenol in patients with severe congestive heart failure. Epoprostenol Multicenter Research Group. *Am J Cardiol* 75: 34A–43A
31. Temple RJ (1995) A regulatory authority's opinion about surrogate endpoints. In: Nimmo WS, Tucker GT (eds) *Clinical measurement in drug evaluation*. John Wiley, New York, pp 3–22
32. UK Prospective Diabetes Study Group (UPDS) (1998) Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet* 352: 837–853
33. Yee KM, Struthers AD (1997) Can drug effects on mortality in heart failure be predicted by any surrogate measure? *Eur Heart J* 18: 1860–1864
34. Yerly S, Pernerger TV, Hirschel B et al. (1998) A critical assessment of the prognostic value of HIV-1 RNA levels and CD4+ cell counts in HIV-infected patients. The Swiss HIV Cohort Study. *Arch Intern Med* 158: 247–252
35. Yusuf S, Hawken S, Ounpuu S et al. (2004) Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 364: 937–952

Hier steht eine Anzeige.

